

Claims:

1. A method for detecting pauses in speech in speech recognition, in which method, for recognizing speech commands uttered by the user, the voice is converted into an electrical signal, **characterized** in that in the method, the frequency spectrum of the electrical signal is divided into two or more sub-bands, samples of the signals in the sub-bands are stored at intervals, the energy levels of the sub-bands are determined on the basis of the stored samples, a power threshold value (thr) is determined, and the energy levels of the sub-bands are compared with said power threshold value (thr), wherein the comparison results are used for producing a pause detecting result.
2. The method according to claim 1, **characterized** in that a detection time limit (END) and a detection quantity (SB_SUFF_TH) are determined, wherein in the method, the calculation of the length of a pause in a sub-band is started when the energy level of the sub-band falls below said power threshold value (thr), wherein in the method, a sub-band specific detection is performed when the calculation reaches the detection time limit (END), it is examined on how many sub-bands the energy level was below the power threshold value (thr) longer than the time detection limit (END), wherein a pause detection decision is made if the number of sub-band specific detections is greater than or equal to the detection quantity (SB_SUFF_TH).
3. The method according to claim 2, **characterized** in that in the method, also an activity time limit (SB_ACTIVE_TH) and an activity quantity (SB_MIN_TH) are determined, wherein a pause detection decision is made if the quantity of sub-band specific detections is greater than or equal to the activity quantity (SB_MIN_TH) and the activity time limit (SB_ACTIVE_TH) has not been reached on the other sub-bands in the calculation of the length of the pause in the sub-band.
4. The method according to claim 1, 2 or 3, **characterized** in that the power threshold value (thr) is calculated by the formula

$$thr = p_{min} + k \cdot (p_{max} - p_{min}), \text{ in which}$$

p_{\min} = the smallest power maximum determined of the stored samples of the sub-bands, and
 p_{\max} = the greatest power minimum determined of the stored samples of the sub-bands.

5

5. The method according to any of the claims 1 to 4, **characterized** in that said power threshold value (thr) is calculated adaptively by taking into account the environmental noise level at each instant.

10

6. The method according to claim 5, **characterized** in that for calculating said power threshold value (thr), a modification coefficient (UPDATE_C) is determined, and on the basis of the stored samples, the greatest power level (win_max) and the smallest power level (win_min) of the sub-bands are calculated, wherein the power maximum (p_{\max}) and power minimum (p_{\min}) are determined by the formulae:

15

$$p_{\max}(i,t) = (1 - \text{UPDATE_C}) \cdot p_{\max}(i,t-1) + (\text{UPDATE_C} \cdot \text{win_max})$$

$$p_{\min}(i,t) = (1 - \text{UPDATE_C}) \cdot p_{\min}(i,t-1) + (\text{UPDATE_C} \cdot \text{win_min})$$

20

in which $0 < \text{UPDATE_C} < 1$,
 $0 < i < L$, and
 L is the number of sub-bands.

25

7. The method according to claim 6, **characterized** in that further in the method,

30

— the modification coefficient (UPDATE_C) is increased, if the absolute value of the difference between said calculated highest power level (win_max) and the power maximum (p_{\max}), or the absolute value of the difference between said calculated lowest power level (win_min) and the power minimum (p_{\min}) has increased,

35

— the modification coefficient (UPDATE_C) is reduced, if the absolute value of the difference between said calculated highest power level (win_max) and the power maximum (p_{\max}), or the absolute value of the difference between said calculated lowest power level (win_min) and the power minimum (p_{\min}) has decreased.

8. A speech recognition device (16) comprising means (1a, 1b) for converting speech commands uttered by a user into an electrical signal, **characterized** in that it also comprises:

- 5 — means (8) for dividing the frequency spectrum of the electrical signal into two or more sub-bands,
- means (14) for storing samples of the signals of the sub-bands at intervals,
- means (5, 13) for determining energy levels of the sub-bands on the basis of the stored samples,
- 10 — means (5, 13) for determining a power threshold value (thr),
- means (5, 13) for comparing the energy levels of the sub-bands with said power threshold value (thr), and
- means (5, 13) for detecting a pause in the speech on the basis of said comparison results.
- 15

9. The speech recognition device (16) according to claim 8, **characterized** in that the power threshold value is calculated by the formula

20 $thr = p_min + k \cdot (p_max - p_min)$, in which

p_min = the smallest determined power maximum of the stored samples of the sub-bands, and

25 p_max = the greatest determined power minimum of the stored samples of the sub-bands.

10. The speech recognition device (16) according to claim 8 or 9, **characterized** in that it comprises also means (10, 11) for filtering the signals of the sub-bands before storage.

30

11. A wireless communication device (MS) comprising means (16) for recognizing speech and means (1a, 1b) for converting speech commands uttered by a user into an electrical signal, **characterized** in that the means (16) for recognizing speech comprise also:

- 35 — means (8) for dividing the frequency spectrum of the electrical signal into two or more sub-bands,
- means (14) for storing samples of the signals of the sub-bands at intervals,

- means (5, 13) for determining energy levels of the sub-bands on the basis of the stored samples,
 - means (5, 13) for determining a power threshold value (thr),
 - means (5, 13) for comparing the energy levels of the sub-bands with said power threshold value (thr), and
 - means (5, 13) for detecting a pause in the speech on the basis of said comparison results.
- 5